UNITÉ DE RECHERCHE
INRIA-ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél.:(1) 39 63 55 11

Rapports de Recherche

N° 1348

Programme 5
Automatique, Productique,
Traitement du Signal et des Données

SENSITIVITY ANALYSIS OF PARAMETRIZED PROGRAMS UNDER CONE CONSTRAINTS

Alexander SHAPIRO
Joseph Frédéric BONNANS

Décembre 1990

RR-1348

ANALYSE DE SENSIBILITE DE PROBLEMES D'OPTIMISATION PARAMETRES SOUS CONTRAINTES DE CONE

SENSITIVITY ANALYSIS OF PARAMETRIZED PROGRAMS UNDER CONE CONSTRAINTS

Alexander Shapiro* Joseph Frédéric Bonnans†

Résumé : Cet article examine le comportement local de solutions de problèmes d'optimisation paramétrés sous contraintes de cônes dans un espace de Banach. Sous certaines conditions suffisantes d'optimalité du deuxième ordre nous établissons la stabilité lipschitzienne des solutions correspondantes. Nous montrons aussi comment ces problèmes d'optimisation paramétriques peuvent être approchées en utilisant le développement au deuxième ordre des données.

Abstract : In this paper we investigate local behavior of optimal solutions of parametrized optimization problems with cone constraints in Banach spaces. Under second-order sufficient optimality conditions we establish Lipschitzian stability of the corresponding optimal solutions. We also show how the considered parametric program can be approximated by using second-order expansions of involved functions.

*Department of Mathematics and Applied Mathematics, University of South Africa, PO Box 392, 0001 Pretoria, South Africa

†INRIA, Domaine de Voluceau, BP 105, 78153 Rocquencourt - France

Abstract

In this paper we investigate local behavior of optimal solutions of parametrized optimization problems with cone constraints in Banach spaces. Under second-order sufficient optimality conditions we establish Lipschitzian stability of the corresponding optimal solutions. We also show how the considered parametric program can be approximated by using second-order expansions of the involved functions.

Key words

Nonlinear optimization, parametric programming, stability and sensitivity analysis, second-order optimality conditions, Lipschitz continuity.

1. Introduction

In this paper we study local behavior of optimal solutions of the parametric optimization problem

$$(\mathscr{P}_t) \qquad \underset{x \in X}{\text{minimize}}\ f(x,t) \quad \text{subject to } x \in \Phi(t),$$

depending on the parameter $t \geq 0$. Here X is a real Banach space, $t \in \mathbb{R}_+$, and it will be supposed throughout the paper that the feasible set $\Phi(t)$ is defined by cone constraints. That is,

$$\Phi(t) = \{x \in X\colon g(x,t) \in K\}, \tag{1.1}$$

where $g:X \times \mathbb{R}_+ \to Y$, Y is a Banach space and K is a closed convex cone in Y. With the program $(\mathscr{P}_t)$ is associated the optimal value function

$$\varphi(t) = \inf\{f(x,t):x \in \Phi(t)\} \tag{1.2}$$

and an ϵ-optimal ($\epsilon \geq 0$) solution $\bar{x}(t)$ satisfying the conditions $x(t) \in \Phi(t)$ and

$$f(\bar{x}(t),t) \leq \varphi(t) + \epsilon. \tag{1.3}$$

The main result of this paper is given in theorem 1 where we show that under certain second-order sufficient conditions and for $\epsilon = \epsilon(t)$ tending to zero sufficiently quickly as $t \to 0^+$, the corresponding ϵ-optimal solutions $\bar{x}(t)$ are upper Lipschitz continuous at t=0. This extends some recent results in sensitivity analysis of nonlinear programs ([4], [5], [7], [13]) to the considered infinite dimensional case and cone constraints. In section 3 we show how the program $(\mathscr{P}_t)$ can be approximated by a simpler one which involves second-order expansions of f(x,t) and g(x,t). This approximation may prove to be useful in calculation of directional derivatives of $\bar{x}(t)$ and requires a further investigation.

We assume that f(x,t) and g(x,t) are *twice Fréchet differentiable*, jointly in x and t, and denote by $D_x f(x,t)$, $D_x g(x,t)$, $D^2_{tt} g(x,t)$, etc, the corresponding partial derivatives. In particular $D^2_{xx} g(x,0)$ belongs to the space $\mathscr{L}(X,\mathscr{L}(X,Y))$ of bounded linear operators from X to the space $\mathscr{L}(X,Y)$, equipped with the corresponding operator norm, and we denote $[D^2_{xx} g(x,0)y]y$ by $D^2_{xx} g(x,0)(y,y)$. We suppose throughout that the null program $(\mathscr{P}_0)$ has a *unique* optimal solution x_0 and that x_0 is a *regular* point of $g(x) = g(x,0)$, with respect to the cone K, in the sense of Robinson [11]. That is

$$0 \in \text{int}\ \{g(x_0) + Dg(x_0)X - K\}. \tag{1.4}$$

The following notation and terminology will be used in the paper. By $c\ell\{S\}$ we denote the topological closure of a set $S \subset X$. For a convex set C and $x \in C$ we denote

$$R(x,C) = \bigcup_{t \geq 0} t(C - x)$$

the radial cone, and $T(x,C) = c\ell\{R(x,C)\}$ the tangent cone to C at x. B(x;r) denotes the ball $\{y \in X: \|y - x\| \leq r\}$ and $B_X = B(0;1)$ denotes the unit ball in X. For $x \in X$ and $\xi \in X^*$ we use the notation $<\xi,x>$ or $<x,\xi>$ for the value $\xi(x)$ of the linear functional ξ at x. If K is a cone in X or in X^*, then its (positive) dual cone K^+ is given by

$$K^+ = \{y: <y,x> \geq 0 \ \text{ for all } x \in K\},$$

and its polar (negative dual) cone is $K^- = -K^+$. By $\partial f(x)$ we denote the subdifferential of a convex function $f(x)$. For a point $x \in X$ and a set $S \subset X$ we denote by $\mathrm{dist}(x,S)$ the distance from x to S.

2. Lipschitz stability of optimal solutions

In this section we study Lipschitz continuity of ϵ-optimal solutions of $(\mathscr{P}_t)$. Note that we assumed existence of the optimal solution for the null (unperturbed) program $(\mathscr{P}_0)$ only. Of course, for $\epsilon > 0$ an ϵ-optimal solution $\bar{x}(t)$ always exists provided that the corresponding feasible set $\Phi(t)$ is nonempty and $\varphi(t) > -\infty$. Sometimes we write $f(x)$, $g(x)$ and Φ for $f(x,0)$, $g(x,0)$ and $\Phi(0)$, respectively.

Under the regularity assumption (1.4) the set

$$\Lambda_0 = \{\lambda \in K^+ : Df(x_0) = \lambda \circ Dg(x_0),\ \langle\lambda, g(x_0)\rangle = 0\}, \tag{2.1}$$

of Lagrange multipliers of the program $(\mathscr{P}_0)$ at the optimal solution point x_0, is nonempty (first-order necessary conditions, [9], [12]) and bounded (e.g. [16, theorem 4.1]). Consequently Λ_0 is a convex and weakly* compact subset of Y^*. Consider the Lagrangian function

$$L(x,\lambda,t) = f(x,t) - \langle\lambda, g(x,t)\rangle$$

of the program $(\mathscr{P}_t)$ and the set

$$\Lambda_1 = \operatorname{argmax}\{D_t L(x_0,\lambda,0) : \lambda \in \Lambda_0\}. \tag{2.2}$$

Notice that the set Λ_1 is nonempty because of the weak* compactness of Λ_0. We shall need the following regularity assumption.

Assumption A.

For some $\lambda_0 \in \Lambda_1$ the tangent cone $T(\lambda_0,\Lambda_0)$ is representable in the form

$$T(\lambda_0,\Lambda_0) = \{\lambda \in T(\lambda_0,K^+) : \lambda \circ Dg(x_0) = 0,\ \langle\lambda, g(x_0)\rangle = 0\}. \tag{2.3}$$

It follows from the definition (2.1) of the set Λ_0 that the radial cone $R(\lambda_0,\Lambda_0)$ can be written in the form

$$R(\lambda_0,\Lambda_0) = \{\lambda \in R(\lambda_0,K^+):\lambda \circ Dg(x_0) = 0, <\lambda,g(x_0)> = 0\}. \tag{2.4}$$

Therefore assumption A holds if and only if the topological closure of the cone given in the right-hand side of (2.4) coincides with the cone given in the right-hand side of (2.3). In particular, assumption A holds in the following cases.

(i) The cone K^+ satisfies at λ_0 the polyhedral property

$$T(\lambda_0,K^+) = R(\lambda_0,K^+). \tag{2.5}$$

(ii) The point x is regular with respect to the cone $K_0 = K(\lambda_0)$, where

$$K(\lambda) = \{y \in K:<\lambda,y> = 0\}. \tag{2.6}$$

Condition (2.5) is relevant to the situations where the feasible set $\Phi(t)$ is defined by equality and a *finite* number of inequality constraints, or when $\lambda_0 = 0$. Condition (ii) was considered in [14]. It implies that the set $\Lambda_0 = \{\lambda_0\}$ is a singleton [14, lemma 4.3] and hence $T(\lambda_0,\Lambda_0) = \{0\}$. Also condition (ii) is equivalent to (see [9, lemma 2.3])

$$Dg(x_0)X - K_0 + [g(x_0)] = Y, \tag{2.7}$$

where $[g(x_0)]$ denotes the linear space generated by vector $g(x_0)$. The polar cone of the cone $Dg(x_0)X - K_0 + [g(x_0)]$ is the intersection of the polar cones of $Dg(x_0)X$, $-K_0$ and $[g(x_0)]$. Since

$$-K_0^- = c\ell(K^+ + [\lambda_0]) = T(\lambda_0,K^+)$$

it follows that the polar cone of $Dg(x_0)X - K_0 + [g(x_0)]$ coincides with the cone given in the right-hand side of (2.3). By (2.7) this polar cone is $\{0\}$ and hence, indeed, condition (ii) implies assumption A.

Lemma 1

Suppose that x_0 *is a regular point of* $g(x)$ *with respect to* K, *that assumption* A *holds and that the cone* $Dg(x_0)X - K_0 + [g(x_0)]$ *is closed. Then there exist positive numbers* κ *and* η *such that*

$$\varphi(t) - \varphi(0) \leq t \max_{\lambda \in \Lambda_0} D_t L(x_0,\lambda,0) + \kappa t^2 \tag{2.8}$$

for all $t \in [0,\eta)$.

Proof

Since λ_0 maximizes $D_t L(x_0,\lambda,0)$ over Λ_0, it follows by the corresponding first-order necessary conditions that

$$D_\lambda[D_t L(x_0,\lambda,0)] = -D_t g(x_0,0) \in N(\lambda_0,\Lambda_0),$$

where $N(\lambda_0,\Lambda_0)$ is the normal cone to Λ_0 at λ_0. The normal cone $N(\lambda_0,\Lambda_0)$ is polar of the cone $T(\lambda_0,\Lambda_0)$ and, because of the assumption A, is given by the topological closure of the cone $Dg(x_0)X - K_0 + [g(x_0)]$. Since it is assumed that the last cone is closed we obtain that

$$-D_t g(x_0,0) \in Dg(x_0)X - K_0 + [g(x_0)].$$

It follows that for any $t > 0$ there exist $\bar{y} \in X$, $k \in K_0$ and $\alpha \in \mathbb{R}$ such that

$$-tD_t g(x_0,0) = tDg(x_0)\bar{y} - k + \alpha g(x_0).$$

Moreover, since $g(x_0) \in K_0$ we can always take $\alpha \geq 0$. Therefore for sufficiently small $t > 0$ we can take $\alpha \in [0,1]$. Then replacing k by $k+(1-\alpha)g(x_0) \in K_0$ we obtain

$$-tD_t g(x_0,0) = tDg(x_0)\bar{y} - k + g(x_0).$$

It follows that

$$g(x_0) + tD_x g(x_0,0)\bar{y} + tD_t g(x_0,0) \in K \tag{2.9}$$

and

$$<\lambda_0, D_x g(x_0,0)\bar{y} + D_t g(x_0,0)> = 0. \tag{2.10}$$

Let us note that since K is convex and $g(x_0) \in K$, if (2.9) holds for some $t=t_0 > 0$ and $\bar{y}$, then it holds for all $t \in [0,t_0]$ and the same $\bar{y}$. Therefore we can choose $\bar{y}$ independently of t for t sufficiently small. Now

$$g(x_0 + t\bar{y},t) = g(x_0) + tD_x g(x_0,0)\bar{y} + tD_t g(x_0,0) + O(t^2). \tag{2.11}$$

Consequently, by the Robinson-Ursescu stability theorem ([11], [15]), it follows from (2.9) and (2.11) that there exists $\tilde{y}(t)$ such that $x_0+\tilde{y}(t) \in \Phi(t)$ and $\|t\bar{y} - \tilde{y}(t)\|$ is of order $O(t^2)$. Then

$$\varphi(t) \leq f(x_0 + \tilde{y}(t),t) = f(x_0) + D_x f(x_0,0)\tilde{y}(t) + tD_t f(x_0,0) + O(t^2)$$

$$= f(x_0) + tD_x f(x_0,0)\bar{y} + tD_t f(x_0,0) + O(t^2).$$

Together with (2.10) this implies

$$\varphi(t) - \varphi(0) \leq tD_x f(x_0,0)\bar{y} - t<\lambda_0, D_x g(x_0,0)\bar{y}> +$$

$$tD_t f(x_0,0) - t<\lambda_0, D_t g(x_0,0)> + O(t^2) =$$

$$t<Df(x_0) - \lambda_0 \circ Dg(x_0)>\bar{y} + tD_t L(x_0,\lambda_0,0) + O(t^2).$$

By the first-order necessary conditions it follows then that

$$\varphi(t) - \varphi(0) \leq tD_t L(x_0,\lambda_0,0) + O(t^2).$$

Since λ_o maximizes $D_t L(x_o,\lambda,0)$ over Λ_o, the inequality (2.8) follows. ■

Remarks

Condition (2.9) holds for some $t > 0$ iff

$$D_x g(x_o,0)\bar{y} + D_t g(x_o,0) \in K + [g(x_o)]. \tag{2.12}$$

Notice that

$$K + [g(x_o)] = R(g(x_o),K).$$

Consider now the following linearizations of the program $(\mathscr{P}_t)$:

$$(\mathscr{L}_i) \qquad \begin{aligned} &\underset{y \in X}{\text{minimize}} \ \langle Df(x_o),y\rangle + D_t f(x_o,0) \\ &\text{subject to } D_x g(x_o,0)y + D_t g(x_o,0) \in M_i, \end{aligned}$$

$i=1,2$, where $M_1 = R(g(x_o),K)$ and $M_2 = T(g(x_o),K)$. Program $(\mathscr{L}_2)$ differs from program $(\mathscr{L}_1)$ only in that its feasible set is the topological closure of the feasible set of program $(\mathscr{L}_1)$. Clearly $\bar{y}$ solves program $(\mathscr{L}_1)$ iff it satisfies condition (2.12) and solves program $(\mathscr{L}_2)$. Therefore a feasible point $\bar{y}$ solves $(\mathscr{L}_1)$ iff there exists $\lambda \in [T(g(x_o),K)]^+$ such that

$$Df(x_o) = \lambda \circ Dg(x_o) \quad \text{and} \quad \langle\lambda, Dg(x_o)\bar{y} + D_t g(x_o,0)\rangle = 0.$$

(Notice that regularity of the program $(\mathscr{L}_2)$ follows from the regularity of the optimal solution x_o of the null program $(\mathscr{P}_o)$.) Since

$$[T(g(x_o),K)]^+ = K^+ \cap \operatorname{Ker} g(x_o),$$

we obtain that $\bar{y}$ solves program $(\mathscr{L}_1)$ iff there exists $\lambda_o \in \Lambda_o$ and $t > 0$ such that conditions (2.9) and (2.10) hold. It follows that under the assumptions of lemma 1, program $(\mathscr{L}_1)$ has a solution.

The dual of programs $(\mathscr{L}_i)$ is the problem of maximization of $D_tL(x_0,\lambda,0)$ subject to $\lambda \in \Lambda_0$. By arguments similar to those of Lempio and Maurer [8, pp, 142-143], it is possible to show that under the assumption of regularity of x_0, *alone*, there is no duality gap between programs $(\mathscr{L}_i)$ and their dual program. Let us briefly outline those arguments.

Consider the optimal value function

$$\psi(v) = \inf\{\langle Df(x_0),y\rangle : Dg(x_0)y + v \in T(g(x_0),K)\}.$$

This is a sublinear (convex and positively homogeneous) function and by the first-order optimality conditions $\psi(0) = 0$. Moreover, it follows from the generalized open mapping theorem [10] that $\psi(v)$ is bounded from above by a finite constant for all v in a neighborhood of zero. This implies that $\psi(v)$ is continuous (e.g. [6, lemma 2.1]) and hence is a support function of a bounded set. We have that $\mu \in \partial\psi(0)$ iff $\psi(v) \geq \langle\mu,v\rangle$ for all $v \in Y$. By duality

$$\psi(v) \geq \max_{\lambda\in\Lambda_0} \langle -\lambda,v\rangle$$

and hence $-\Lambda_0 \subset \partial\psi(0)$. Also, by taking y=0 in the definition of $\psi(v)$ we obtain that for $\mu \in \partial\psi(0)$,

$$\langle\mu,v\rangle \leq 0 \quad \text{for all } v \in T(g(x_0),K),$$

and hence

$$-\mu \in T(g(x_0),K)^+ = K^+ \cap \operatorname{Ker} g(x_0).$$

Furthermore, for a given y taking $v = -Dg(x_0)y$, we obtain

$$<Df(x_0),y> + <\mu,Dg(x_0)y> \geq 0$$

and hence

$$Df(x_0) + \mu \circ Dg(x_0) = 0.$$

Consequently, $-\mu \in \Lambda_0$ and hence $\partial\psi(0) = -\Lambda_0$. It follows that

$$\psi(v) = \max_{\lambda \in \Lambda_0} <-\lambda,v>.$$

The fact that there is no duality gap between program $(\mathscr{L}_1)$ and its dual, is sufficient to show that the inequality (2.8) holds but with the term κt^2 is replaced by $o(t)$ (see [8, theorem 3.1]).

We employ the following second-order sufficient condition. For $\eta \geq 0$ consider the cone

$$C_\eta = \{y \in X{:}Dg(x_0)y \in K + [g(x_0)],\ <Df(x_0),y> \leq \eta\|y\|\} \tag{2.13}$$

and the set Λ_1 defined in (2.2).

<u>Assumption B</u> (*second-order sufficient condition*)
There exist $\alpha > 0$ and $\eta > 0$ such that

$$\max_{\lambda \in \Lambda_1} <y,D^2_{xx}L(x_0,\lambda,0)y> \geq \alpha\|y\|^2 \tag{2.14}$$

for all $y \in C_\eta$.

For η=0 the corresponding cone C_0 is called the critical cone of the program $(\mathscr{P}_0)$. Notice that by the first-order necessary conditions, $<Df(x_0),y>$ is nonnegative for all y such that $Dg(x_0)y$ belongs to the radial cone $R(g(x_0),K)$ (e.g., [9, theorem 3.1]). Therefore the second-term inequality in the right-hand side definition

(2.13) of C_0 can be replaced by the equation $\langle Df(x_0),y\rangle = 0$. The second-order condition of assumption B is a natural extension of the (strong) second-order sufficient condition employed in [13, p. 635] for finite dimensional cases and a finite number of constraints.

Assumption B implies that for any positive number β less than α, there exists a neighborhood W of x_0 such that

$$f(x) \geq f(x_0) + \beta\|x - x_0\|^2$$

for all $x \in \Phi \cap W$ (cf. [9, theorem 5.6]). It follows then that if $\bar{x}(t)$ is an $\epsilon(t)$-optimal solution of $(\mathscr{P}_t)$, $\epsilon(t) = O(t)$ and $\bar{x}(t) \in W$, then $\bar{x}(t)$ converges to x_0 as $t \to 0^+$ at least at a rate of $O(t^{1/2})$ ([1], [2], [3], [14]). In the following theorem we establish Lipschitzian rate of convergence of $\bar{x}(t)$ to x_0.

Theorem 1

Let $\epsilon(t) = O(t^2)$ *and* $\bar{x}(t)$ *be an* $\epsilon(t)$*-optimal solution of* $(\mathscr{P}_t)$ *converging to* x_0 *as* $t \to 0^+$. *Suppose that the assumptions of lemma* 1 *and assumption* B *hold. Then there exists a positive constant* c *such that*

$$\|\bar{x}(t) - x_0\| \leq ct \tag{2.15}$$

for all $t \geq 0$ *sufficiently small.*

Proof

Suppose that (2.15) is false. Then there are $t_n \to 0^+$, $x_n = \bar{x}(t_n)$ and $\tau_n = \|x_n - x_0\|$ such that

$$\lim_{n\to\infty} t_n/\tau_n = 0. \tag{2.16}$$

Let η be a positive constant specified in assumption B and consider $y_n = \tau_n^{-1}(x_n - x_0)$. We have by (2.16) that

$$g(x_n,t_n) = g(x_o) + \tau_n Dg(x_o)y_n + o(\tau_n).$$

Since $g(x_n,t_n) \in K$ it follows then that

$$\mathrm{dist}(Dg(x_o)y_n, K + [g(x_o)]) \to 0.$$

By the generalized open mapping theorem (or the Robinson-Ursescu stability theorem) this implies that

$$\mathrm{dist}(y_n,S) \to 0$$

where

$$S = \{y \in X:Dg(x_o)y \in K + [g(x_o)]\}.$$

Therefore there exist $\bar{y}_n \in S$ such that $\|y_n - \bar{y}_n\|$ tends to zero as $n \to \infty$. Now by the definition of ϵ-optimality

$$\varphi(t_n) - \varphi(0) \geq f(x_n,t_n) - f(x_o,0) - \epsilon_n,$$

where $\epsilon_n = \epsilon(t_n)$. It follows then by (2.16) that

$$\varphi(t_n) - \varphi(0) \geq \tau_n \langle y_n, Df(x_o)\rangle + o(\tau_n).$$

Moreover, because of (2.8),

$$\limsup_{n\to\infty} \tau_n^{-1}[\varphi(t_n) - \varphi(0)] \leq 0.$$

It follows that

$$\langle \bar{y}_n, Df(x_o)\rangle \leq \eta,$$

for n large enough, and hence $\bar{y}_n \in C_\eta$.

Now since for every $\lambda \in \Lambda_0$,

$$\varphi(0) = L(x_0,\lambda,0)$$

and

$$\varphi(t_n) \geq L(x_n,\lambda,t_n) - \epsilon_n$$

we have that

$$\varphi(t_n) - \varphi(0) \geq \max_{\lambda\in\Lambda_0} \{L(x_n,\lambda,t_n) - L(x_0,\lambda,0)\} - \epsilon_n.$$

Furthermore,

$$L(x_n,\lambda,t_n) - L(x_0,\lambda,0) = t_n D_t L(x_0,\lambda,0) + \tfrac{1}{2}\tau_n^2 \langle y_n, D_{xx}^2 L(x_n^*,\lambda,t_n^*) y_n\rangle$$

$$+ t_n \tau_n \langle y_n, D_{xt}^2 L(x_n^*,\lambda,t_n^*)\rangle + \tfrac{1}{2} t_n^2 D_{tt}^2 L(x_n^*,\lambda,t_n^*),$$

where (x_n^*,t_n^*) is a point on the segment joining $(x_0,0)$ and (x_n,t_n). It follows then by continuity of the second-order derivatives, boundedness of Λ_0 and (2.16) that

$$\varphi(t_n) - \varphi(0) \geq \max_{\lambda\in\Lambda_0} \{t_n D_t L(x_0,\lambda,0) + \tfrac{1}{2}\tau_n^2 \langle y_n, D_{xx}^2 L(x_0,\lambda,0) y_n\rangle\} + o(\tau_n^2)$$

$$\geq t_n \max_{\lambda\in\Lambda_0} \{D_t L(x_0,\lambda,0)\} + \tfrac{1}{2}\tau_n^2 \max_{\lambda\in\Lambda_1} \{\langle y_n, D_{xx}^2 L(x_0,\lambda,0) y_n\rangle\} + o(\tau_n^2).$$

Since $\|y_n - \bar{y}_n\|$ tends to zero and $\bar{y}_n \in C_\eta$ for n large enough we have by the second-order condition of assumption B that

$$\max_{\lambda\in\Lambda_1} \langle y_n, D_{xx}^2 L(x_0,\lambda,0) y_n\rangle \geq \tfrac{1}{2}\alpha$$

and hence

$$\varphi(t_n) - \varphi(0) \geq t_n \max_{\lambda\in\Lambda_0} \{D_t L(x_0,\lambda,0)\} + \tfrac{1}{4}\alpha\tau_n^2 + o(\tau_n^2).$$

The last inequality contradicts the result (2.8) of lemma 1 and hence the proof is complete. ∎

Lipschitzian stability of optimal solutions of parametrized programs was studied in a recent paper of Alt [2]. His result [2, theorem 3.4] requires a regularity condition involving Lagrange multipliers which may be difficult to verify in situations where the set Λ_0 is not a singleton.

As we mentioned earlier, if the point x_0 is regular with respect to the cone $K_0 = K(\lambda_0)$, given in (2.6), then $\Lambda_0 = \{\lambda_0\}$ and assumption A holds. In this case Lipschitzian stability of optimal solutions of $(\mathscr{P}_t)$ implies Lipschitzian stability of the corresponding Lagrange multipliers [14, lemma 4.4]. It is interesting to note that regularity of x_0 with respect to the cone K, uniqueness of λ_0 and Lipschitzian stability of the optimal solutions do not imply Lipschitzian stability of the corresponding Lagrange multipliers even in the finite dimensional case. We show this in the following example.

<u>Example</u>
Let $X = \mathbb{R}^2$, $Y = \mathbb{R}^3$ and consider

$$f(x) = x_1 + x_2 + x_1^2 + x_2^2,$$

$g(x) = Gx$, where G is the 3×2 matrix

$$G^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and the cone

$$K = \{y : y_3^2 \geq y_1^2 + y_2^2,\ y_3 \geq 0\}.$$

Notice that $K^+ = K$. Then $x_0 = (0,0)$ is the optimal solution of the problem,

$$\text{minimize } f(x) \text{ subject to } g(x) \in K.$$

The corresponding first-order necessary conditions, are

$$(\lambda_1, 0) + (0, \lambda_3) = (1,1), \quad \lambda = (\lambda_1, \lambda_2, \lambda_3) \in K^+,$$

which give the unique Lagrange multiplier $\lambda_0 = (1,0,1)$. It can be easily verified that the point $x_0 = (0,0)$ is a regular point of $g(x)$ with respect to the cone K but not with respect to the cone K_0.

Consider the following perturbations of the objective function

$$f(x,t) = (1 - t)x_1 + (1 + t)x_2 + x_1^2 + x_2^2.$$

Then for all $t \in [0,1)$, $x_0 = (0,0)$ is the optimal solution with the corresponding set of Lagrange multipliers

$$\Lambda(t) = \{(1-t,\lambda_2,1+t):|\lambda_2| \le 2t^{1/2}\}.$$

In particular $\bar{\lambda}(t) = (1-t,2t^{1/2},1+t)$ is a vector of the Lagrange multipliers with $\|\bar{\lambda}(t) - \lambda_0\| \ge t^{1/2}$.

3. Second-order expansion of $(\mathscr{P}_t)$

In this section we show how second-order expansions of $f(x,t)$ and $g(x,t)$ can be employed in order to approximate the program $(\mathscr{P}_t)$ by a simpler one. Consider $a_0 = Df(x_0)$, $G_0 = Dg(x_0)$, $c_0 = g(x_0)$ and for $\eta \ge 0$ the set

$$\Lambda_\eta = \bigcup_{(a,G,c)\in\Omega_\eta} \{\lambda \in K^+:a = \lambda\circ G, \langle\lambda,c\rangle = 0\}, \tag{3.1}$$

where

$$\Omega_\eta = \{(a,G,c):\|a - a_0\| + \|G - G_0\| + \|c - c_0\| \le \eta\}.$$

Clearly, Λ_η contains the set Λ_0 of Lagrange multipliers and for $\eta = 0$ both sets coincide. Note that under the assumption (1.4) of regularity of x_0, the set Λ_η is bounded for sufficiently small $\eta > 0$. Indeed, consider $\lambda \in \Lambda_\eta$ and for a given $\epsilon > 0$ let $y \in B_Y$ be such that $\langle\lambda,y\rangle \ge \|\lambda\| - \epsilon$. Then by the generalized open mapping theorem, it follows from the regularity of x_0 that there are a positive constant α, $x \in B_X$, $|\gamma| \leq 1$ and $k \in K$ such that

$$y = \alpha(G_0 x + \gamma c_0) - k$$

and $\alpha \leq \bar{\alpha}$, where $\bar{\alpha}$ does not depend on y.

Let $(a,G,c) \in \Omega_\eta$ be such that $a = \lambda \circ G$ and $<\lambda,c> = 0$. We have then that

$$<\lambda,y> \leq \alpha<\lambda,G_0x+\gamma c_0> = \alpha<\lambda,Gx+\gamma c> + \alpha<\lambda,(G_0-G)x + \gamma(c_0-c)>$$

$$\leq \alpha\|a\| + \alpha\|\lambda\|(\|G_0-G\| + \|c_0-c\|) \leq \alpha\|a_0\| + \alpha\eta + \alpha\eta\|\lambda\|.$$

It follows that for $\eta < \bar{\alpha}^{-1}$,

$$\|\lambda\| \leq (1 - \bar{\alpha}\eta)^{-1} (\bar{\alpha}\|a_0\| + \bar{\alpha}\eta + \epsilon)$$

which shows that Λ_η is bounded.

We shall need the following strong form of second-order sufficient conditions.

Assumption C

There exist $\beta > 0$ and $\eta > 0$ such that for any $\lambda \in \Lambda_\eta$,

$$<y,D^2_{xx}L(x_0,\lambda,0)y> \geq \beta\|y\|^2 \tag{3.2}$$

for all $y \in X$.

Note that assumption C implies existence of a quadratic form, on the space X, which induces a norm equivalent to the original norm $\|\cdot\|$. Consequently X becomes a Hilbert space.

Now let us consider the program,

$$(\mathcal{Q}_t) \qquad \begin{array}{l} \underset{y \in X}{\text{minimize}} \ \langle y, D_x f(x_o,t)\rangle + \frac{1}{2}\langle y, D^2 f(x_o) y\rangle \\ \\ \text{subject to } g(x_o,t) + D_x g(x_o,t)y + \frac{1}{2} D^2_{xx} g(x_o,0)(y,y) \in K. \end{array}$$

Notice that under the assumptions of theorem 1, it follows from the Lipschitz stability of $\bar{x}(t)$ that the optimal value function $\varphi(t)$ is differentiable at t=0 (in the positive direction) and

$$\varphi'(0) = \max_{\lambda \in \Lambda_o} D_t L(x_o, \lambda, 0)$$

(cf. [8, theorem 3.4]). Similar result holds for the program $(\mathcal{Q}_t)$ as well and hence the difference between the optimal value functions of programs $(\mathcal{P}_t)$ and $(\mathcal{Q}_t)$ is of order o(t).

Theorem 2

Suppose that the assumptions of lemma 1 and assumption C hold and that for all sufficiently small $t \geq 0$ *program* $(\mathcal{Q}_t)$ *has an optimal solution* $y^*(t)$. *Let, for* $\epsilon(t) = o(t^2)$, $\bar{x}(t)$ *be an* $\epsilon(t)$*-optimal solution of* $(\mathcal{P}_t)$ *converging to* x_o *as* $t \to 0^+$. *Then*

$$\|\bar{x}(t) - x_o - y^*(t)\| = o(t). \tag{3.3}$$

Proof

Consider the functions

$$f^*(y,t) = \langle y, D_x f(x_o,t)\rangle + \frac{1}{2}\langle y, D^2 f(x_o) y\rangle,$$

$$g^*(y,t) = g(x_o,t) + D_x g(x_o,t)y + \frac{1}{2} D^2_{xx} g(x_o,0)(y,y),$$

the Lagrangian

$$L^*(y,\lambda,t) = f^*(y,t) - \langle \lambda, g^*(y,t)\rangle$$

and the feasible set

$$\Psi(t) = \{y \in X : g^*(y,t) \in K\}$$

corresponding to the program $(\mathcal{Q}_t)$. We have that for any $\lambda \in \Lambda_0$ and $y \in \Psi(0)$,

$$f^*(y,0) \geq L^*(y,\lambda,0).$$

Since for any $\lambda \in \Lambda_0$,

$$L^*(y,\lambda,0) = \tfrac{1}{2}\langle y, D^2_{xx}L(x_0,\lambda,0)y\rangle,$$

it follows then from assumption C that

$$f^*(y,0) \geq \tfrac{1}{2}\beta\|y\|^2 \tag{3.4}$$

for all $y \in \Psi(0)$, and hence $y=0$ is the optimal solution of $(\mathcal{Q}_0)$. Since

$$\|D_y f^*(y,t) - D_y f^*(y,0)\| = \|D_x f(x_0,t) - D_x f(x_0,0)\| \to 0$$

as $t \to 0^+$, it also follows from (3.4) and regularity of x_0 that $y^*(t) \to 0$ as $t \to 0^+$. By theorem 1 we have then that $\|\bar{x}(t) - x_0\|$ and $\|y^*(t)\|$ are of order $O(t)$. Therefore there is $c > 0$ such that $\|\bar{x}(t) - x_0\| \leq ct$ and $\|y^*(t)\| \leq ct$ for all sufficiently small $t \geq 0$.

By continuity of $D_x f(x,t)$ at $(x_0,0)$ we have that $f(\cdot,t)$ and $f^*(\cdot,t)$ are Lipschitz continuous in neighborhoods of x_0 and 0, respectively, with Lipschitz constant independent of t for all t small enough. Also, by the Robinson-Ursescu stability theorem

$$\delta_1(t) = \sup_{x \in (x_0+\Psi(t)) \cap B(x_0;ct)} \operatorname{dist}(x,\Phi(t)) = o(t^2)$$

and

$$\delta_2(t) = \sup_{y\in(\Phi(t)+x_0)\cap B(0;ct)} \operatorname{dist}(y,\Psi(t)) = o(t^2).$$

Furthermore, consider

$$\kappa(t) = \sup_{\|x-x_0\|\leq ct} \|D_x f(x,t) - D_x f^*(x-x_0,t)\|$$

$$= \sup_{\|x-x_0\|\leq ct} \|D_x f(x,t) - D_x f(x_0,t) - D^2 f(x_0,0)(x - x_0)\|,$$

which is of order $o(t)$ because of second-order Fréchet differentiability of $f(x,t)$. Now, for a given $t \geq 0$ small enough, consider $y^* = y^*(t)$. Because of the regularity of x_0 and since $y^*(t) \to 0$ as $t \to 0^+$, we have that there is a Lagrange multiplier $\lambda^* = \lambda^*(t)$ corresponding to the optimal solution $y^*(t)$ of the program $(\mathcal{Q}_t)$ such that $\lambda^* \in \Lambda_\eta$ for sufficiently small t. It follows that for all $y \in \Psi(t)$,

$$f^*(y,t) - f^*(y^*,t) \geq L^*(y,\lambda^*,t) - L^*(y^*,\lambda^*,t). \tag{3.5}$$

Since the right-hand side of (3.5) is equal to $\frac{1}{2}\langle y-y^*, D^2_{xx}L(x_0,\lambda^*,0)(y-y^*)\rangle$ we obtain by assumption C that

$$f^*(y,t) - f^*(y^*,t) \geq \tfrac{1}{2}\beta\|y-y^*\|^2$$

for all $y \in \Psi(t)$ and all t small enough. Now (3.3) follows by lemma 2.2 in [14] and hence the proof is complete. ■

Remarks

In the case

$$\lim_{\eta\to 0^+} \{ \sup_{\lambda\in\Lambda_\eta} \operatorname{dist}(\lambda,\Lambda_0)\} = 0, \tag{3.6}$$

the set Λ_η in assumption C can be replaced by Λ_0. Condition (3.6) holds, at least, in the following cases:

(i) The linear space generated by the cone K^+ is finite-dimensional.

(ii) The point x_0 is regular with respect to the cone $K_0 = K(\lambda_0)$, $\Lambda = \{\lambda_0\}$.

Program $(\mathscr{Q}_t)$ can be expanded around $t_0 = 0$ as well. That is, we can consider the program

$$\underset{y\in X}{\text{minimize}} <y,Df(x_0) + tD^2_{xt}f(x_0,0)> + \tfrac{1}{2}<y,D^2f(x_0)y>$$

$(\mathscr{Q}'_t)$

$$\text{subject to } g(x_0) + tD_tg(x_0,0) + D_xg(x_0,0)y +$$

$$\tfrac{1}{2}t^2D^2_{tt}g(x_0,0) + tD^2_{xt}g(x_0,0)y + \tfrac{1}{2}D^2_{xx}g(x_0,0)(y,y) \in K.$$

Under assumptions similar to those of theorem 2, an optimal solution $y'(t)$ of $(\mathscr{Q}'_t)$ will provide a first-order approximation of $\bar{x}(t) - x_0$.

Existence of the optimal solution $y^*(t)$ (optimal solution $y'(t)$) follows from assumption C if the program $(\mathscr{Q}_t)$ (program $(\mathscr{Q}'_t)$) is convex. In particular, if $D^2_{xx}g(x_0,0) = 0$ (for example, if $g(x,0)$ is linear in x), then $D^2_{xx}L(x_0,\lambda,0)$ is equal to $D^2_{xx}f(x_0,0)$ for every λ and the constraint mappings of the programs $(\mathscr{Q}_t)$ and $(\mathscr{Q}'_t)$ are linear in y. Therefore in this case existence of $y^*(t)$ and $y'(t)$ is guaranteed by assumption C.

References

[1] Alt, W., Lipschitzian perturbations of infinite optimization problems. *Mathematical Programming with Data Perturbations II*, Edited by A.V. Fiacco, New York, pp. 7-21, 1983.

[2] Alt, W., Stability of solutions for a class of nonlinear cone constrained optimization problems, Part 1: basic theory. *Numerical Functional Analysis and Optimization*, 10 (1989), 1053-1064.

[3] Auslender, A., Stability in mathematical programming with nondifferentiable data. *SIAM J. Control and Optimization*, 22 (1984), 239-254.

[4] Auslender, A and Cominetti, R., First and second-order sensitivity analysis of nonlinear programs under directional constraint qualification conditions. *Optimization*, 21 (1990), 351-363.

[5] Bonnans, J.F., Directional derivatives of optimal solutions in smooth nonlinear programming. *Journal Optimization Theory and Application*, to appear.

[6] Ekeland, I. and Temam, R., *Convex Analysis and Variational Problems.* North-Holland, Amsterdam, 1976.

[7] Gauvin, J. and Janin, R., Directional behavior of optimal solutions in nonlinear mathematical programming. *Mathematics of Operations Research*, 13 (1988), 629-649.

[8] Lempio, F. and Maurer, H., Differential stability in infinite-dimensional nonlinear programming. *Applied Mathematics and Optimization*, 6 (1980), 139-152.

[9] Maurer, H. and Zowe, J., First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems. *Mathematical Programming* 16 (1979), 98-110.

[10] Robinson, S.M., Normed convex processes. *Transactions of the American Mathematical Society*, 174 (1972), 127-140.

[11] Robinson, S.M., Stability theorems for systems of inequalities, Part II: differentiable nonlinear systems. *SIAM J. Numerical Analysis*, 13 (1976), 497-513.

[12] Robinson, S.M., First-order conditions for general nonlinear optimization. *SIAM J. Applied Mathematics*, 30 (1976), 597-607.

[13] Shapiro, A., Sensitivity analysis of nonlinear programs and differentiability properties of metric projections. *SIAM J. Control and Optimization*, 26 (1988), 628-645.

[14] Shapiro, A., Perturbation analysis of optimization problems in Banach spaces. Preprint, University of South Africa, 1990.

[15] Ursescu, C., Multifunctions with convex closed graph. *Czechoslovak Mathematical Journal*, 25 (1975), 438-441.

[16] Zowe, J. and Kurcyusz, S., Regularity and stability for the mathematical programming problem in Banach spaces. *Applied Mathematics and Optimization*, 5 (1979), 49-62.

Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique.

ISSN 0249-6399